A family of mammalian F-box proteins

Jeffrey T. Winston*, Deanna M. Koepp*†‡, Cihui Zhu*, Stephen J. Elledge*†‡ and J. Wade Harper*

Ubiquitin-mediated destruction of regulatory proteins is a frequent means of controlling progression through signaling pathways [1]. F-box proteins [2] are components of modular E3 ubiquitin protein ligases called SCFs, which function in phosphorylation-dependent ubiquitination ([3–5], reviewed in [6,7]). F-box proteins contain a carboxy-terminal domain that interacts with substrates and a 42–48 amino-acid F-box motif which binds to the protein Skp1 [2–4]. Skp1 binding links the F-box protein with a core ubiquitin ligase composed of the proteins Cdc53/Cul1, Rbx1 (also called Hrt1 and Roc1) and the E2 ubiquitin-conjugating enzyme Cdc34 [8–11]. The genomes of the budding yeast *Saccharomyces cerevisiae* and the nematode worm *Caenorhabditis elegans* contain, respectively, 16 and more than 60 F-box proteins [2,7]; in *S. cerevisiae*, the F-box proteins Cdc4, Grr1 and Met30 target cyclin-dependent kinase inhibitors, G1 cyclins and transcriptional regulators for ubiquitination ([3–5,8,10], reviewed in [6,7]). Only four mammalian F-box proteins (Cyclin F, Skp1, β-TRCP and NFB42) have been identified so far [2,12]. Here, we report the identification of a family of 33 novel mammalian F-box proteins. The large number of these proteins in mammals suggests that the SCF system controls a correspondingly large number of regulatory pathways in vertebrates. Four of these proteins contain a novel conserved motif, the F-box-associated (FBA) domain, which may represent a new protein–protein interaction motif. The identification of these genes will help uncover pathways controlled by ubiquitin-mediated proteolysis in mammals.

Addresses: *Verna and Marrs McLean Department of Biochemistry and Molecular Biology, †Howard Hughes Medical Institute, and ‡Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA.

Correspondence: J. Wade Harper
E-mail: jharper@bcm.tmc.edu

Received: 12 August 1999
Revised: 13 September 1999
Accepted: 13 September 1999

Published: 11 October 1999

Current Biology 1999, 9:1180–1182

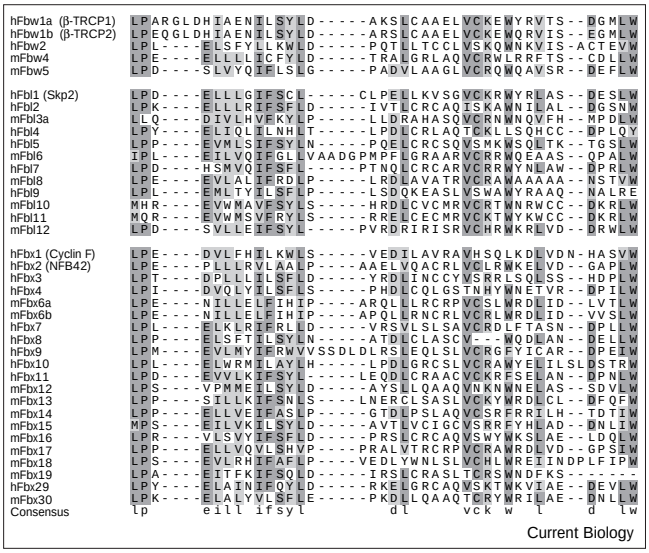
0960-9822/99/$ – see front matter

Results and discussion

Using yeast two-hybrid screens to isolate proteins that bind Skp1 and by searching the expressed sequence tag (EST) database, we identified 33 cDNAs encoding novel mammalian F-box proteins (Figure 1). In 24 cases, both human and rodent homologs were identified, and they were typically > 90% identical (data not shown). Cenciarelli *et al.* [13] and Regan-Reimann *et al.* [14] have independently identified a partially overlapping set of human and *Xenopus* F-box proteins. A composite alignment of 50 mammalian F-box motifs is included in the Supplementary material.

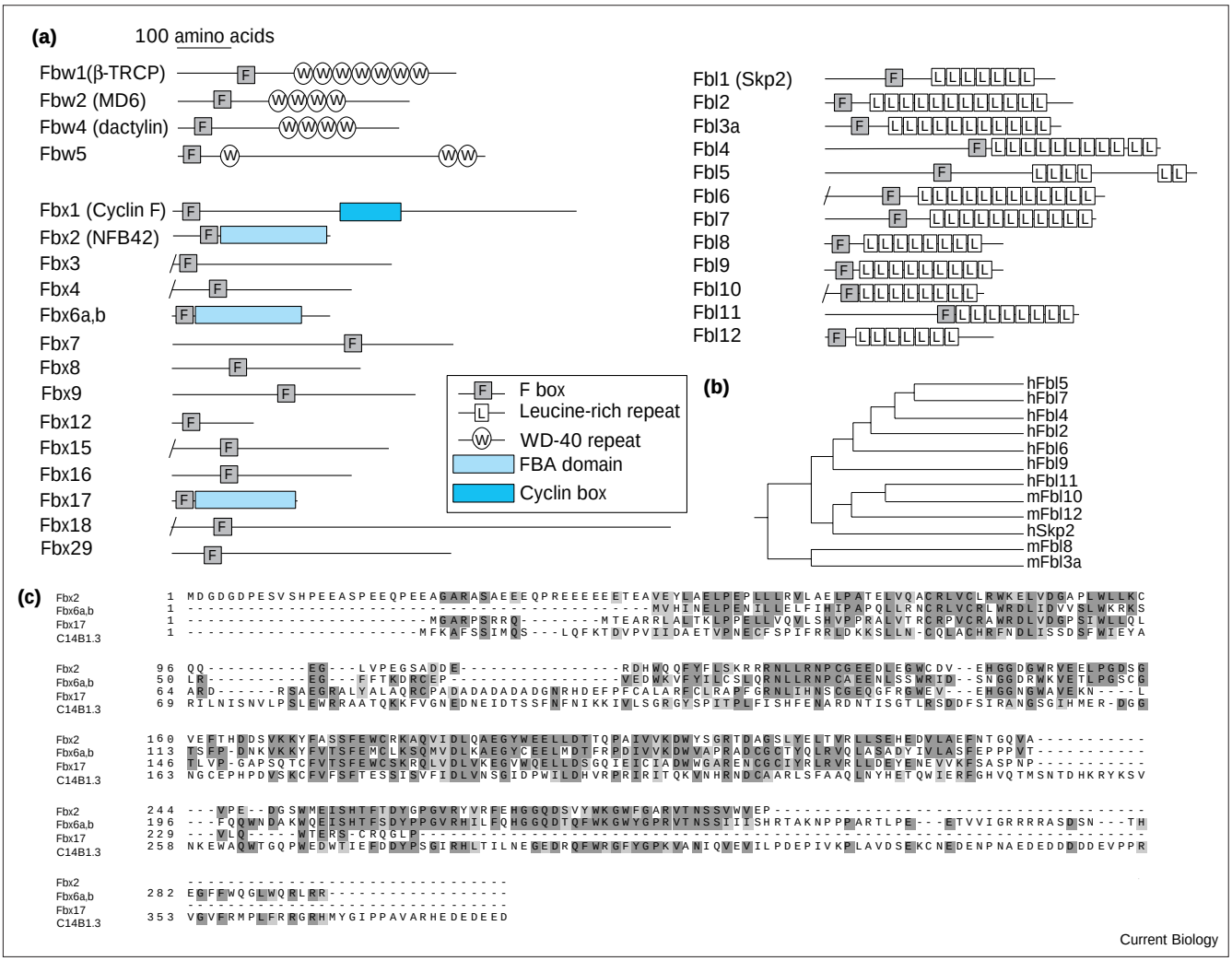
Sequence analysis revealed three subfamilies of F-box proteins: Fbws contain WD-40 repeats, Fbls contain leucine-rich repeats, and Fbxs lack known protein-interaction domains (Figures 1,2a); β-TRCP, Skp2 and Cyclin F, respectively, are the first members of the three subfamilies, and the two close homologs of β-TRCP will be referred to as Fbw1a and Fbw1b. In agreement with Cenciarelli *et al.* [13], the other members have been numbered in order of discovery.

Figure 1



A family of mammalian F-box proteins. Sequences of novel F-box motifs and comparison with F-box motifs in Cyclin F (Fbx1), Skp2 (Fbl1), β-TRCP1 (Fbw1a), β-TRCP2 (Fbw1b) and NFB42 (Fbx2). The alignment was performed using Clustal W 1.7. Dark gray, identical residues; light gray, similar residues; h, human; m, mouse. Fbx3 was identified from a human breast cDNA library using Skp1 as bait in a yeast two-hybrid screen. Other novel F-box sequences were obtained by performing iterative BLAST searches of the GenBank database using sequences of newly-identified F-box motifs. GenBank accession numbers are provided in the Supplementary material or can be obtained by querying GenBank at http://www.ncbi.nlm.nih.gov/genbank/query_form.html.

Figure 2



(a) Domain structures of F-box proteins. (b) Phylogenetic analysis of Fbl proteins, performed using the DNAStar software package. (c) Alignment of those members of the Fbx subfamily that have FBA domains: Fbx2, Fbx6a, Fbx17 and the hypothetical *C. elegans* protein C14B1.3. Dark gray, identical residues; light gray, similar residues

The Fbls comprise the largest subfamily of F-box proteins, with 12 members (Figure 2a) that, on the basis of phylogenetic comparisons, cluster into three groups (Figure 2b). Fbl2, Fbl5 and Fbl7 display extensive internal homology within the leucine-rich repeats and are most closely related to Grr1, which is involved in G1 cyclin ubiquitination in budding yeast [3,6,8,10], and the hypothetical F-box protein C02F5.7 in *C. elegans*. Other Fbl proteins display little internal similarity in their leucine-rich repeats or with one another outside of the LXL motifs (in the single-letter amino-acid code, where X is any amino acid) that define this domain. Fbw2 and Fbw4 contain multiple WD-40 repeats but display no significant sequence similarities with each other or with other proteins in the database outside these motifs (Figure 2a). Fbw5 contains three WD-40 repeats, the first one separated by a large spacer region from the other two. Recently, mutations in Fbw4 were shown to be responsible for dactylaplasia in mice, a condition that resembles split-hand split-foot malformation-3 in humans [15].

The Fbx subfamily are a diverse set of proteins, lacking defined protein–protein interaction domains carboxy-terminal to the F box, with the exception of the cyclin box found in Cyclin F (Figure 2a). Nevertheless, four members displayed similarity to each other and were characterized by a carboxy-terminal motif, which we call the F-box-associated (FBA) domain, possibly representing a new protein–protein interaction motif (Figure 2a,c). The founding member of this group is Fbx2 (NFB42), recently

reported to be abundant in neurons [12]. We found three additional mammalian homologs of Fbx2: the two closely related proteins Fbx6a and Fbx6b and the more distantly related Fbx17 (Figure 2a,c). *C. elegans* has two hypothetical F-box proteins (C14B1.3 and T01E8.4), which have FBA-like domains (Figure 2c and data not shown).

The F-box motif is functionally defined as a motif that can interact with Skp1 [2]. We found that 12 randomly selected F-box proteins from all three subfamilies (Fbw2, Fbw5, Fbl2, Fbl4, Fbl5, Fbl8, Fbx3, Fbx7, Fbx8, Fbx12, Fbx15, Fbx16) associate with a fusion protein between glutathione-S-transferase and Skp1 (GST–Skp1) *in vitro* (see Supplementary material) and with Skp1 in transfected 293 cells (data not shown). These 12 proteins are widely expressed during mouse embryogenesis and in adult tissues, and preliminary studies indicate that the majority are located in the cytoplasm when expressed transiently in 293 cells (see Supplementary material).

In summary, mammalian F-box proteins represent an expanding family of proteins and additional members are likely to emerge with time. Currently, F-box proteins are the largest class of E3 ubiquitin ligase receptors and, as many F-box proteins recognize multiple substrates, the SCF system may be able to ubiquitinate hundreds of proteins. Some F-box proteins may themselves be targeted for ubiquitination through association with the Skp1–Cul1 complex [2,16]. Recent data suggest that other homologs of cullin proteins may be involved in ubiquitin ligase systems [9,11] that are also combinatorial in nature. For example, the Cul2–Elongin BC complex interacts with a family of SOCS-box proteins (typified by the Suppressor of cytokine signalling-1 protein), of which there are currently 20 members [17,18]. The identification of mammalian F-box proteins will facilitate both the elucidation of pathways that are controlled by the SCF system and the identification of particular F-box proteins that recognize known ubiquitination targets. For example, surveys of a collection of F-box proteins for interaction with ubiquitination substrates has led to the identification of β-TRCP (Fbw1a) as the mediator of IκBα and β-catenin ubiquitination [19], and Skp2 (Fbl1) as the mediator of p27 ubiquitination [20]. Identification of targets of the family of mammalian F-box proteins reported here will be a challenge for the future.

Supplementary material

Supplementary material including figures showing alignment of leucine-rich repeats, association of F-box proteins with Skp1, subcellular localization and expression patterns; GenBank accession numbers; primer sequences; and a compilation of F-box sequences from this paper and [13] is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We thank M. Pagano and P. Jackson for communicating results before publication. This work was supported by grants from the NIH (AG-11085) to J.W.H. and S.J.E., the Welch Foundation (Q-1249) to J.W.H., the Department of Defense to J.W.H, and the Baylor SPORE in Prostate Cancer. J.T.W. was supported by a postdoctoral training grant from NIA, and D.M.K. was supported by the Helen Hay Whitney Foundation. S.J.E. is an investigator with the Howard Hughes Medical Institute.

References

1. Hershko A, Ciechanover A: The ubiquitin system. *Annu Rev Biochem* 1998, 67:425-479.
2. Bai C, Sen P, Mathias N, Hofmann K, Ma L, Goebl M, Harper JW, Elledge SJ: SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* 1996, 86:263-274.
3. Skowyra D, Craig K, Tyers M, Elledge SJ, Harper JW: F-box proteins are components of E3 complexes and act as receptors to recruit phosphorylated substrates for ubiquitination. *Cell* 1997, 91:209-219.
4. Feldman RM, Correll CC, Kaplan KB, Deshaies RJ: A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* 1997, 91:221-230.
5. Skowyra D, Koepp DM, Kamura T, Conrad MN, Conaway RC, Conaway JW, Elledge SJ, Harper JW: Reconstitution of G1 cyclin ubiquitination with complexes containing SCF^{Grr1} and Rbx1. *Science* 1999, 284:662-665.
6. Elledge SJ, Harper JW: Proteolysis in cell cycle control and cancer. *Biochem Biophys Acta* 1998, 1377:M61-M70.
7. Patton EE, Willems AR, Tyers M: Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. *Trends Genet* 1998, 14:236-243.
8. Patton EE, Willems AR, Sa D, Kuras L, Thomas D, Craig KL, Tyers M: Cdc53 is a scaffold protein for multiple Cdc34/Skp1/F-box protein complexes that regulate cell division and methionine biosynthesis in yeast. *Genes Dev* 1998, 12:692-705.
9. Kamura T, Koepp DM, Conrad MN, Skowyra D, Moreland RJ, Iliopoulos O, *et al.*: Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science* 1999, 284:657-661.
10. Seol JH, Feldman RM, Zachariae W, Shevchenko A, Correll CC, Lyapina S, *et al.*: Cdc53/cullin and the essential hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme cdc34. *Genes Dev* 1999, 13:1614-1626.
11. Ohta T, Michel JJ, Schottelius AJ, Xiong Y: ROC1, a homolog of APC11, represents a family of cullin partners with an associated ubiquitin ligase activity. *Mol Cell* 1999, 3:535-541.
12. Erhardt JA, Hynicka W, DiBenedetto A, Shen N, Stone N, Paulson H, Pittman RN: A novel F box protein, NFB42, is highly enriched in neurons and induces growth arrest. *J Biol Chem* 1998, 273:35222-35227.
13. Cenciarelli C, Chiaur DS, Guardavaccaro D, Parks W, Vidal M, Pagano M: Identification of a family of human F-box proteins. *Curr Biol* 1999, 9:1177-1179.
14. Regan-Reimann JD, Duong QV, Jackson PK: Identification of Skp1 interacting F-box proteins in ***Xenopus laevis***. *Curr Biol* 1999, 9:R762-R763.
15. Sidow A, Bulotsky MS, Kerrebrock AW, Birren BW, Altshuler D, Jaenisch R, *et al.*: A novel member of the F-box/WD40 gene family, encoding dactylin, is disrupted in the mouse dactylaplasia mutant. *Nat Genet* 1999, 23:104-107.
16. Zhou P, Howley PM: Ubiquitination and degradation of the substrate recognition units of SCF ubiquitin-protein ligases. *Mol Cell* 1998, 2:571-580.
17. Kamura T, Sato S, Haque D, Liu L, Kaelin WG Jr, Conaway RC, Conaway JW: The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev* 1998, 12:3872-3881.
18. Hilton DJ, Richardson RT, Alexander WS, Viney EM, Willson TA, Sprigg NS, *et al.*: Twenty proteins containing a C-terminal SOCS box form five structural classes. *Proc Natl Acad Sci USA* 1998, 95:114-119.
19. Winston JT, Strack P, Beer-Romero P, Chu CY, Elledge SJ, Harper JW: The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination ***in vitro***. *Genes Dev* 1999, 13:270-283.
20. Carrano AC, Eytan E, Hershko A, Pagano M: Skp2 is required for the ubiquitin-mediated degradation of the CDK inhibitor p27. *Nat Cell Biol* 1999, 1:193-199.

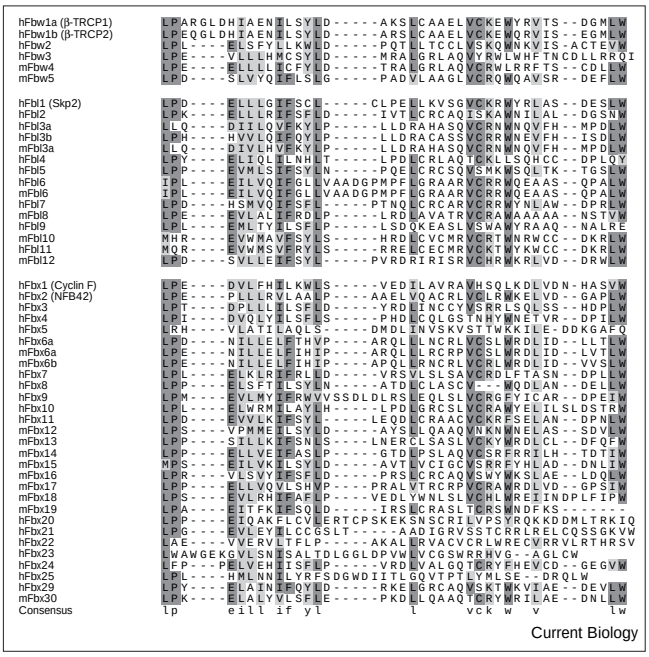
Supplementary material

A family of mammalian F-box proteins

Jeffrey T. Winston, Deanna M. Koepp, Cihui Zhu, Stephen J. Elledge and J. Wade Harper

Current Biology 11 October 1999, 9:1180–1182

Figure S1



Composite alignment of F-box proteins. The sequences of F-box motifs reported in the paper and those by Cenciarelli *et al.* [S1] were compiled and aligned using Clustal 1.7 followed by minor visual realignment. Dark gray, identical residues; light gray, similar residues; h, human; m, mouse.

Supplementary materials and methods

Accession numbers of F-box proteins

hFbw1a, AF110396; hFbw1b, BAA31671; hFbw2, AF176698; mFbw4, AF176519; mFbw5, AF176520; hFbl1, U33761; hFbl2, AF176518; mFbl3a, AF176521; hFbl4, AF176699; hFbl5, AF176700; mFbl6, AF176522; hFbl7, AB020647; mFbl8, AF176523; hFbl9, AF176701; mFbl10, AF176524; hFbl11, AB023221; mFbl11, Al154332; mFbl12, AF176525; hFbx1, U17105; hFbx2, AF187318; hFbx3, AF176702; hFbx4, AF176703; mFbx6a, AU067142; mFbx6b, AF176526; hFbx7, AA315010, AL050254.1; mFbx8, AF176527; hFbx9, AF176704, AL031178.1; hFbx10, AF176705; hFbx11, AF176706; mFbx12, AF176528; mFbx13, AF176529; mFbx14, AU066822; mFbx15, AF176530; mFbx16, AF176531; mFbx17, AF176532; mFbx18, AF184275; mFbx19, AA501293; hFbx29, AF176707; mFbx30, Al836688.

Primer sequences

hFb12: forward, 5′-TGCAGCTAGGACCATGCCAGAAACCTG-3′; reverse, 5′-GAAGACCCTGCTCCACCCAGTATAGAG-3′; hFb14: forward, 5′-GAGCAGCACCGGGTGCTTCACCAGA-3′; reverse, 5′-GAGAGCGCAGCTCATGTAAATTCTCAC-3′; hFb15: forward, 5′-CCAGTGAATTGTAATGTTGTTTCTCT-3′; reverse, 5′-TGTGGGTT-TGTTTGAATTGGCCTTCTG-3′; hFb17: forward, 5′-GAGGCA-CCAGCACACCTAATCACAGCAAAG-3′; reverse, 5′-GTCTAATCA TCTCAGTGTGTTCCCTCTTC-3′; hFb19: forward, 5′-GTCAGT-AACCAGCCCTGCAGCACAG-3′; reverse, 5′-AGAGTCCCACAG-GAGTAGGTGGAGC-3′; hFbx4: forward, 5′-CCAAGCACC-CAG-GGAATTCTGATGCATTCC-3′; reverse, 5′-GCAACTGCTTTC-ATGC-CTAACAGATCACAG-3′; hFbx7: forward, 5′-GTGATGCTCCTGC-CATCGTCAACTCAC-3′; reverse, 5′-GTCACAGAGGAACCT-TTG-GGTCATACC-3′; hFbx10: forward, 5′-GCCCAGCTTCTAGCTCCC-CAAAGCCAGGCT-3′; reverse, 5′-CACCGCCCTCCAAGCGAAGA-AAGATGCCTG-3′; mFbw2: forward, 5′-AGGAAATCAAAGAGAG-GCACCAGCTTCCTG-3′; reverse, 5′-GCATGCTGCGGTGTATGTC-AGTAAACAAGC-3′; mFbw5: forward, 5′-GAGCCATTGGAACC-TACTGGGGAAAATGGT-3′; reverse, 5′-CACAGTCCACTGCAGGC-CACTCAGTAG-3′; mFblB: forward, 5′-TGCGCGGGCTTGAGA-GAAATACACTGCTTC-3′; reverse, 5′-CAGCTCTGGACCTGAGCTA-GAGCTAGTGCG-3′; mFbxB: forward, 5′-CATCTACCTTATTGGC-CATGTGGCTGC-3′; reverse, 5′-GGATAGTTCCCAGACCAATCTAA-TGCAC-3′; mFbx15: forward, 5′-CCTGTCAAACGCCGCACAAGC-CTTCCTTCG-3′; reverse, 5′-GCCCATTATGCTTCTCTCGAAAAG-GCCATG-3′; mFbx16: forward, 5′-GAAGTGACCCCAGACTTC-AAGCGCCAG-3′; reverse, 5′-GACTAGCTAGAATCCAACTC-TCAGCTG-3′.

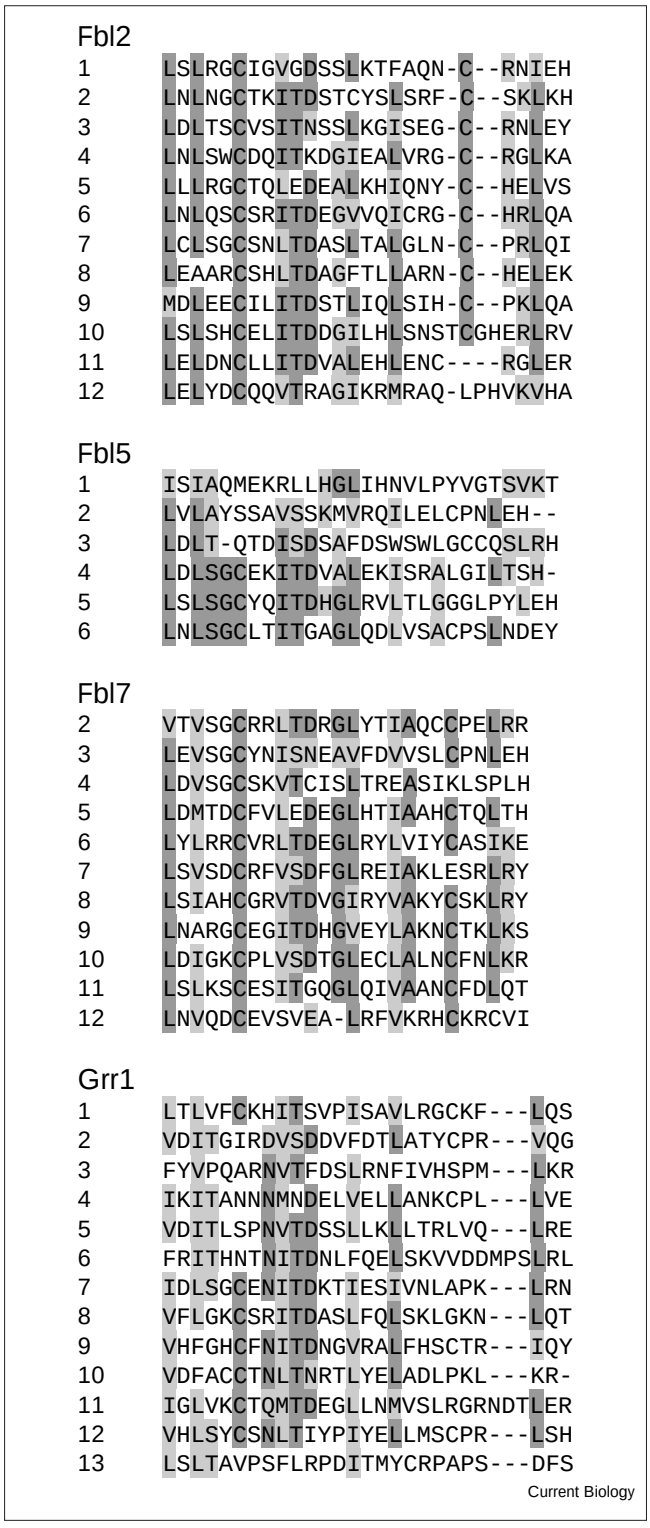
Supplementary references

S1. Cenciarelli C, Chiaur DS, Guardavaccaro D, Parks W, Vidal M, Pagano M: Identification of a family of human F-box proteins. *Curr Biol* 1999, 9:1177-1179.

S2. Liu Q, Li MZ, Leibham D, Cortez D, Elledge SJ: The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr Biol* 1998, 8:1300-1309.

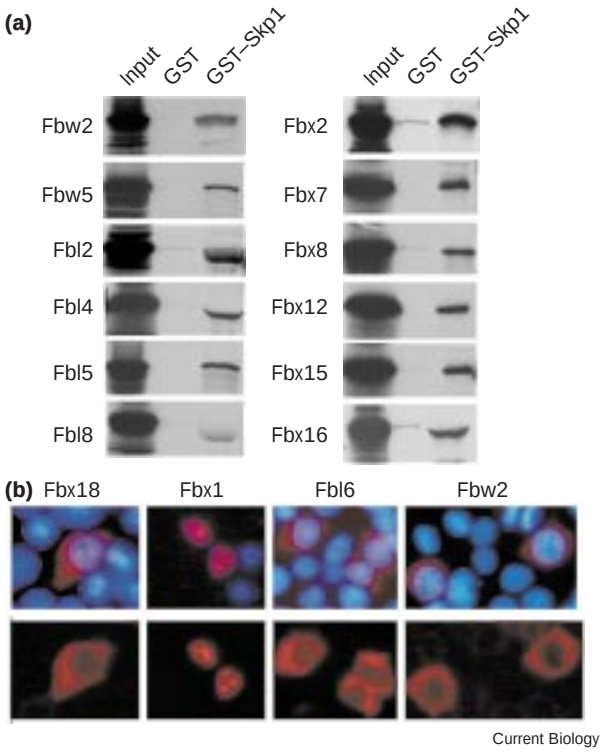
S3. Bai C, Richman R, Elledge SJ: Human Cyclin F. *EMBO J* 1994, 13:6087-6098.

Figure S2



Alignment of leucine-rich repeat elements in several F-box proteins. The sequences were aligned using Clustal 1.7. The Cys residue at position 6 and the Ile–Thr–Asp segment beginning at position 9 are highly conserved in Fbl2, Fbl5, and Fbl7 and are characteristic of the yeast F-box protein Grr1.

Figure S3



Association of Skp1 with F-box proteins. (**a**) GST–Skp1 and GST were purified from insect cells using glutathione–Sepharose and 0.5 μg immobilized protein used for *in vitro* binding with 5 μl *in vitro* translated and [^{35}S]methionine-labeled hemagglutinin (HA)-tagged F-box protein. The cDNAs for F-box proteins were cloned into the Univector pUNI10 [S2] and Cre-mediated plasmid fusion used to place cDNAs under CMV/T7 control while simultaneously fusing coding sequences to HA epitope tag. Vectors were used for *in vitro* translation using the TNT system (Promega). After washing with buffers containing 0.5% NP40, proteins were separated by SDS–PAGE and visualized by autoradiography. Input corresponds to 100% of that used for *in vitro* binding. (**b**) Subcellular localization of F-box proteins. Cytomegalovirus (CMV) expression vectors expressing HA-tagged F-box proteins were transfected into 293 cells using lipofectin and, after 36 h, cells were subjected to immunofluorescence using anti-HA antibodies and detection with Texas Red-conjugated secondary antibody. Cells were stained with 4′,6-diamidino-2-phenylindole (DAPI; blue) to visualize nuclei. The majority of F-box proteins examined were localized predominantly in the cytoplasm, with a small fraction of cells (< 20%) displaying both nuclear and cytoplasmic localization (representative examples are shown). With Fbl4, Fbx3, and Fbx7, a larger percentage of the cells (60–80%) displayed both nuclear and cytoplasmic staining, with the remainder displaying exclusively cytoplasmic localization. Consistent with previous reports [S3], we found that HA-tagged Cyclin F (Fbx1), used here as a control, was largely nuclear. Similar results were observed using COS cells (data not shown).

Table S1

Expression of mRNAs of F-box proteins in tissues.

cDNA	Brain	Heart	Kidney	Liver	Lung	Skeletal muscle	Pancreas	Placenta	Spleen	Testis	Day 7 embryo	Day 11 embryo	Day 15 embryo	Day 17 embryo
hFbl2	+	+	+	+	+	–	+	+						
hFbl4	ND	+	+	+	+	–	+	+						
hFbl5	+	++	+	+	+	+	++	+						
hFbl7	+	++	++	++	++	+	+	+						
hFbl9	–	+	+	+	+	–	+	–						
hFbx4	–	+	+	+	+	–	+	+						
hFbx7	++	++	++	++	++	++	+	++						
hFbx10	+	+	+	+	+	+	+	+						
mFbw2	+	+	+	+	+	+			+	+	+	+	+	+
mFbw5	++	++	++	++	++	++			+	++	++	+	++	++
mFbl8	+	+	+	+	++	+			+	+	++	–	+	+
mFbx8	++	++	++	++	++	++			+	++	++	+	+	+
mFbx15	–	–	–	–	–	–			–	++	–	+	–	–
mFbx16	+	+	+	–	+	–			–	+	–	+	+	+

Expression of mRNAs of F-box proteins was determined by quantitative PCR unless otherwise noted. A panel of cDNAs from the indicated tissues (Clontech) was used for amplification as recommended by the supplier. Primer sets were designed such that PCR products were in the 300–500 bp range; for primer sequences, see Supplementary materials and methods. Reactions were performed in a total of 50 μl with a 30 sec denaturation (94°C) followed by 22–38 cycles of denaturation (94°C) and annealing/elongation (68°C), depending on the abundance of individual F-box protein, and aliquots from various cycles analyzed to ensure that amplification was in the linear range. PCR products for each primer set were directly sequenced to verify that they represent the appropriate cDNA. ND, not determined; –, no product detectable with 38 cycles of PCR; +, PCR product readily detectable with 30 cycles; ++, PCR product readily detectable with 22 cycles.